2020

Overexpression of ferulate 5-hydroxylase increases syringyl units in *Sorghum bicolor*

Hannah M. Tetreault
USDA, Agricultural Research Service, hannah.tetreault@ars.usda.gov

Tammy Gries
USDA ARS, tammy.gries@usda.gov

Nathan Palmer
USDA ARS, nathan.palmer@ars.usda.gov

Deanna L. Funnell-Harris
University of Nebraska-Lincoln, Deanna.Funnell-Harris@ars.usda.gov

Shirley Sato
University of Nebraska - Lincoln, ssato1@unl.edu

See next page for additional authors

Plant Molecular Biology (2020) 103:269–285
https://doi.org/10.1007/s11103-020-00991-3

Overexpression of ferulate 5-hydroxylase increases syringyl units in *Sorghum bicolor*

Hannah M. Tetreault[1,2] · Tammy Gries[1] · Nathan A. Palmer[1] · Deanna L. Funnell-Harris[1,3] · Shirley Sato[2,4] · Zhengxiang Ge[2,4] · Gautam Sarath[1,2] · Scott E. Sattler[1,2]

Received: 21 June 2019 / Accepted: 4 March 2020 / Published online: 13 March 2020

Abstract

Ferulate 5-hydroxylase (F5H) of the monolignol pathway catalyzes the hydroxylation of coniferyl alcohol, coniferaldehyde and ferulic acid to produce 5-hydroxyconiferyl moieties, which lead to the formation of sinapic acid and syringyl (S) lignin monomers. In contrast, guaiacyl (G) lignin, the other major type of lignin monomer, is derived from polymerization of coniferyl alcohol. In this study, the effects of manipulating S-lignin biosynthesis in sorghum (*Sorghum bicolor*) were evaluated. Overexpression of sorghum *F5H* (*SbF5H*), under the control of the CaMV *35S* promoter, increased both S-lignin levels and the ratio of S/G lignin, while plant growth and development remained relatively unaffected. Maüle staining of stalk and leaf midrib sections from *SbF5H* overexpression lines indicated that the lignin composition was altered. Ectopic expression of *SbF5H* did not affect the gene expression of other monolignol pathway genes. In addition, *brown midrib 12-ref* (*bmr12-ref*), a nonsense mutation in the sorghum caffeic acid *O*-methyltransferase (COMT) was combined with *35S::SbF5H* through cross-pollination to examine effects on lignin synthesis. The stover composition from *bmr12 35S::SbF5H* plants more closely resembled *bmr12* stover than *35S::SbF5H* or wild-type (WT) stover; S-lignin and total lignin concentrations were decreased relative to WT or *35S::SbF5H*. Likewise, expression of upstream monolignol biosynthetic genes was increased in both *bmr12* and *bmr12 35S::SbF5H* relative to WT or *35S::SbF5H*. Overall, these results indicated that overexpression of *SbF5H* did not compensate for the loss of COMT activity.

Key message **Overexpression of F5H in sorghum increases S-lignin without increasing total lignin content or affecting plant growth, but it cannot compensate for the loss of COMT activity in monolignol synthesis.**

Keywords *Sorghum bicolor* · Ferulate 5-hydroxylase (F5H) · *Brown midrib 12* (*bmr12*) · Caffeic acid *O*-methyltransferase · Syringyl lignin (s-lignin)

Introduction

The heterogeneous phenolic polymer lignin is found in secondary plant cell walls, cross-linking with cell wall polysaccharides to form an intricate, rigid matrix. Lignin is critical for structural integrity in vascular plants providing support and water transport as well as plant defense (Boerjan et al. 2003). However, considerable research on lignin biosynthesis has focused on decreasing lignin content in biomass because it is the major factor in the recalcitrance of renewable lignocellulosic biomass to efficient processing for a range of fuel or chemical applications (Chen and Dixon 2007; Dien et al. 2009; Ragauskas et al. 2014). Three major monolignols, *p*-coumaryl, coniferyl and sinapyl alcohol form the lignin polymer through radical coupled polymerization which give rise to *p*-hydroxyphenol (H), guaiacyl (G) and

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s11103-020-00991-3) contains supplementary material, which is available to authorized users.

✉ Scott E. Sattler
scott.sattler@usda.gov

[1] Wheat, Sorghum and Forage Research Unit, USDA-ARS, Lincoln, NE 68583, USA

[2] Department of Agronomy and Horticulture, University of Nebraska-Lincoln, Lincoln, NE 68583, USA

[3] Department of Plant Pathology, University of Nebraska-Lincoln, Lincoln, NE 68583, USA

[4] Center for Plant Science Innovation, University of Nebraska, Lincoln, NE 68588, USA

syringyl (S) units, respectively. The proportions of the three major monolignols vary among taxa, species and even cell types (Chapple et al. 1992). In addition, the composition of these monomers within cell walls is flexible, which allows researchers to manipulate steps in lignin synthesis that result in changes to lignin composition and content. Monolignols and other phenylpropanoid compounds are incorporated into a lignin polymer through oxidative radical coupling that results in a heterogenous polymer whose structure and pattern of constituents may be highly variable (Hatfield et al. 2017). Coniferyl alcohol forms more types of cross-coupling reactions than sinapyl alcohol upon radicalization due to the absence of the 5-methoxy from its phenol ring (Ralph et al. 2004). Thus, G-lignin is more condensed, more highly cross-linked and less linear than S-lignin, which may affect the conversion of biomass to sugars and aromatic compounds (Guo et al. 2001).

The C4 grass sorghum (*Sorghum bicolor*), which is native to Africa and able to be grown across a broad range of environments with minimal agronomic inputs (Rooney 2004), is being developed as a lignocellulosic feedstock (Sarath et al. 2008; Mitchell et al. 2016; Lee et al. 2018). Reduced lignin concentration has been shown to increase forage digestibility and increase conversion efficiency for cellulosic ethanol using *brown midrib* (*bmr*) mutants of sorghum (Cherney et al. 1991; Oliver et al. 2005; Dien et al. 2009; Sattler et al. 2012; Godin et al. 2016). In sorghum and other C4 grasses, *bmr* mutants have a distinguishable tan to reddish brown coloration present in their leaf midribs compared to green or white midribs of wild-type (WT) plants, and this visible phenotype has long been associated with impaired ability to synthesize lignin (Jorgenson 1931; Porter et al. 1978).

The *bmr* mutants have been used to identify and characterize the genes encoding three main enzymes of monolignol biosynthesis in sorghum (Fig. 1). The sorghum *Bmr2* gene encodes a 4-coumarate coenzyme-A ligase (4CL), which catalyzes the initial steps in monolignol biosynthesis that results in the formation of 4-coumaroyl CoA (Saballos et al. 2012). The sorghum *Bmr6* gene encodes a cinnamyl alcohol dehydrogenase (CAD; Saballos et al. 2009; Sattler et al. 2009), which catalyzes the conversion of monolignals to monolignols. Sorghum *Bmr12* encodes caffeic acid *O*-methyltransferase (COMT), which catalyzes the conversion of 5-hydroxy-coniferylaldehyde to sinapylaldehyde (Bout and Vermerris 2003). Thus, the *bmr* mutants are valuable tools to study monolignol biosynthesis in sorghum and provide opportunities to manipulate the concentration and composition of lignin within biomass for bioenergy uses.

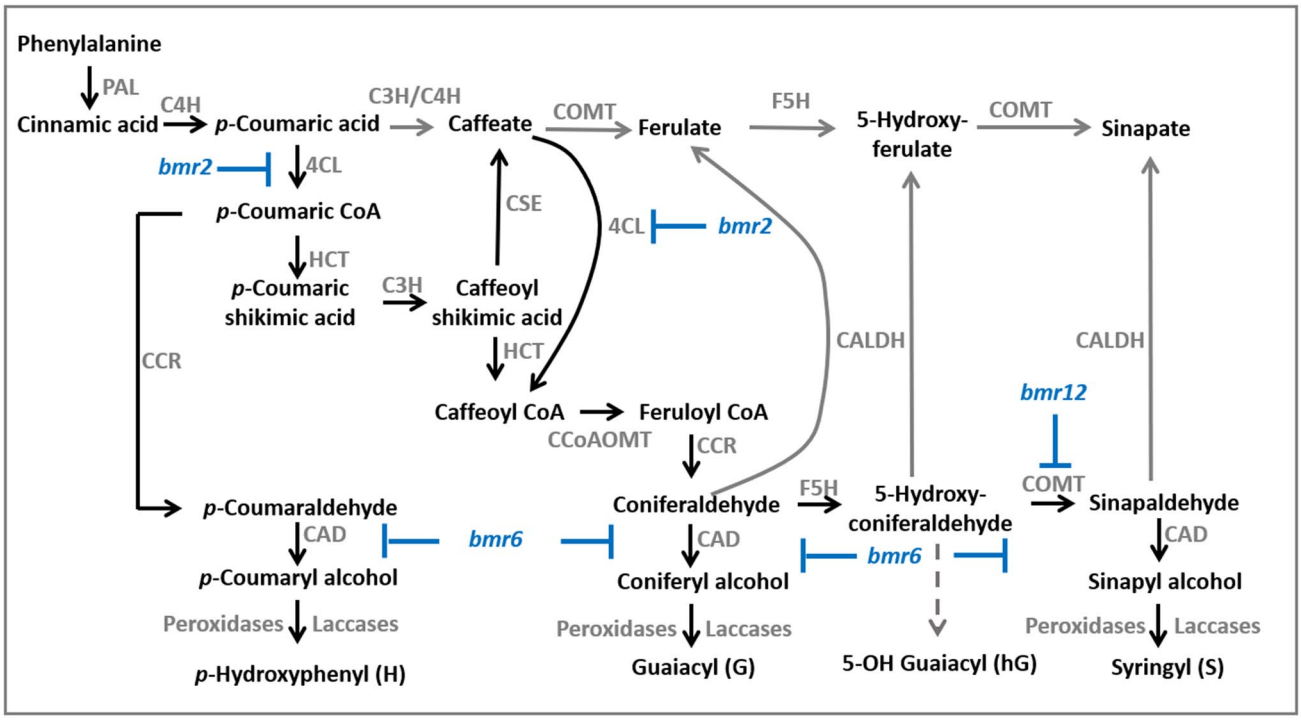


Fig. 1 The monolignol biosynthesis pathway in sorghum based on consensus model from dicot and monocot plants. Enzymes along the pathway (gray) represent: *PAL* phenylalanine ammonia lyase, *C4H* cinnamate 4-hydroxylase, *4CL* 4-coumarate-CoA ligase, *HCT* *p*-hydroxycinnamoyltransferase, *C3′H* *p*-coumaroyl quinate/shikimate 3′-hydroxylase, *CSE* caffeoyl shikimate esterase, *CCoAOMT* caffeoyl-CoA-*O*-methyltransferase, *CCR* cinnamoyl-CoA reductase, *F5H* ferulate 5-hydroxylase, *COMT* caffeic acid *O*-methyltransferase, *CALDH* cinnamyl aldehyde dehydrogenase, *CAD* cinnamyl alcohol dehydrogenase. Gray lines indicate proposed steps in the pathway. The *bmr2*, *bmr6* and *bmr12* mutants and enzymes impaired are indicated in blue

Two enzymes are responsible for converting coniferyl moieties into sinapyl moieties and influencing the ratio of guaiacyl to syringyl lignin (G/S) within cell walls. The ferulate 5-hydroxylase (F5H), a cytochrome P450-dependent monooxygenase, hydroxylates the phenyl ring of coniferylaldehyde to form 5-hydroxyconiferylaldehyde (Humphreys et al. 1999) and a COMT (Bmr12) methylates the 5-hydroxyl group of 5-hydroxyconiferylaldehyde to form sinapylaldehyde (Fig. 1). Thus, both F5H and COMT have been a focus for altering lignin composition of vascular plants. The *fah1* Arabidopsis mutant, deficient in F5H activity, has almost no S-lignin, while overexpression of *F5H* increases the S-unit content of lignin (Meyer et al. 1998; Franke et al. 2000; Stewart et al. 2009). Reduction of COMT activity through either mutation or co-suppression reduced S-lignin and resulted in the incorporation of 5-hydroxyconiferyl alcohol to form 5-hydroxyguaiacyl (5-OH-G) lignin subunits in sorghum, maize, poplars and Arabidopsis (Vignols et al. 1995; Jouanin et al. 2000; Ralph et al. 2001; Rastogi and Dwivedi 2006; Palmer et al. 2008). Overexpression of *F5H* in the COMT-deficient Arabidopsis mutant, *omt1* resulted in lignin polymers dramatically reduced in G and S units as well as substantially increased in 5-OH-G-lignin units, however growth and development were negatively impacted (Weng et al. 2010). Similarly, overexpression of *F5H* in COMT-silenced switchgrass reduced G-lignin as well as increased 5-OH-G-lignin units (Wu et al. 2019). Overall, these results showed that manipulating monolignol biosynthesis can markedly alter lignin composition, which may lead to changes in cell wall digestibility and plant growth.

In the present study, sorghum F5H (*SbF5H*; Sobic.001G196300.1) was overexpressed in sorghum using a constitutive promoter to characterize its impact on cell wall composition and plant fitness. In addition, *SbF5H* overexpression lines were crossed with *bmr12* plants and the effects of these combined alterations on lignin composition and content of sorghum biomass were assessed.

Methods

Generation of transgenic *SbF5H* overexpression lines

The coding region of *S. bicolor* ferulate-5-hydroxylase (Sb*F5H*; Sobic.001G196300.1) was amplified by PCR with the primers SbF5H_Pcil-F, 5′-CGTACATGTCGGCCGTTG CCAAGATCGCC-3′ and SbF5H_Xbal-R 5′-TTTTCTAGA TCAGTACAGGGGGCAGTTGAGC-3′, using Turbo *Pfu* polymerase (Agilent) and the University of Georgia EST clone PH1_9_G09_A002 (GenBank accession CF428636) as the template. The coding region was subcloned between the E35S CaMV promoter and the 35S CaMV terminator as a PciI and XbaI fragment and submitted to Eurofins Genomics (https://www.eurofinsgenomics.com) for DNA sequencing to confirm DNA sequence fidelity. The *E35S::SbF5H* cassette in the pZP211 binary vector was transformed into grain *S. bicolor* (RTx430) using *Agrobacterium tumefaciens.* Seven independent transgenic events were generated and *F5H* expression (T3 generation) quantified via RT-qPCR following methods in Scully et al. (2016). Two homozygous transgenic lines (ZG251-1-16a and ZG251-1-19a), referred to as SbF5H-16a and SbF5H-19a were selected based on *F5H* expression levels and crossed with *bmr12* (RTx430) (Pedersen et al. 2006). NN and ZG are abbreviations for the technical staff that conducted transformation experiments.

Plant materials and growth conditions

Seeds from the *SbF5H-16a/bmr12* and *SbF5H-19a/bmr12* crosses were planted and seedlings screened via PCR for the four genotypic classes: WT, *35S::SbF5H*, *bmr12* and stacked (*bmr12 35S::SbF5H*), genotypic classes are referred to as WT(16a/19a), *bmr12*(16a/19a), *SbF5H*-16a/19a and stacked(16a/19a). Briefly, genomic DNA was extracted separately from leaf tissue of each seedling using a cetyltrimethyl ammonium bromide (CTAB) method (Rogers and Bendich 1985). Presence of the *SbF5H* transgene was confirmed using PCR amplification with *F5H* detection primers (forward primer: 5′-AGATGAGCTCGCCCAGGTTCA-3′ and reverse primer 5′-TTCATTTGGAGAGGACCTCGAGA-3′). Amplifications were carried out using a Biometra T gradient PCR System with an initial denaturation step at 95 °C for 5 min followed by 40 cycles of 95 °C for 30 s, 60 °C for 30 s and 72 °C for 45 s and a final extension step at 72 °C for 5 min. *Bmr12* genotype was determined using high resolution melt curve analysis. PCR amplification was performed using primers (forward primer: 5′-AGTGGCTCACCCCTA ACGAG-3′ and reverse primer 5′-TGTGCTCTGACGACG ACTAC-3′) and Precision Melt Supermix (Bio-Rad) with a Bio-Rad CFX96 Touch Real-Time PCR Detection System according to manufacturer's instructions with the following parameters: initial denaturation step at 98 °C for 2 min, followed by 40 cycles at 98 °C for 5 s, 60 °C for 5 s and a final melting step from 60 to 95 °C with an increase of 0.2 °C every 5 s. Melt curve analysis was performed by collecting data from the melting step and results were analyzed with the Bio-Rad Melt Curve Analysis Software. Seedlings identified as homozygous for each genotypic class were then grown to maturity and seed was collected for further characterization. Seeds for each genotypic class were planted in a soil mixture with a 1:2:1:1 ratio of soil:peat moss:vermiculite:sand. All plants were grown under a 16:8 h light:dark cycle in a randomized complete block design in the University of Nebraska-Lincoln Greenhouse Facility. Watering was conducted daily or as needed and fertilization with a weak

nutrient solution (Dyna Green All Purpose 12–12–12) was applied weekly. At 5–6 weeks after germination, the fifth leaf from the base and 10 cm of stalk tissue were harvested for RNA extraction, immediately flash-frozen in liquid nitrogen, ground using a freezer mill (SPEX SamplePrep) and stored at − 80 °C. Leaf midrib and stalk materials were also collected for microscopy at this stage. Leaf midrib tissue was collected at the 3 cm section closest to the stalk of the fifth leaf and stalk tissue was collected from the top internode under the peduncle. The remaining plants were grown to maturity for stover (biomass) analyses (fiber, bomb calorimetry, thioacidolysis and phenolics analyses). Panicles were separated from stover, and all tissues were dried in forced-air ovens at 50 °C. Dried stover, which encompassed both leaf and stalk tissue, was ground in a Wiley mill fitted with a 2-mm mesh screen (Arthur H. Thomas Co.), followed by grinding on a cyclone mill fitted with a 1-mm mesh screen (UDY Co.).

RNA extraction and quantitative RT-PCR

Total RNA was extracted from leaf and stalk tissue from three individual plants per genotypic class. Approximately 100 mg of homogenized plant material was added to 1 ml of TriPure Isolation Reagent (Sigma-Aldrich) followed by RNA extraction and purification using the RNA Clean and Concentrator Kit (Zymo Research). RNA was treated with an on-column DNase treatment (Zymo Research). RNA integrity was confirmed using a 1.8% denaturing agarose gel stained with ethidium bromide (EtBr). RNA quantity was determined using a Synergy Microplate (BioTek Instruments). One microgram of total RNA was utilized to generate a cDNA library for real-time qPCR using the Evagreen chemistry on a Fluidigm Biomark HD Instrument (Fluidigm) following manufacturer supplied protocols (available on-line at www.fluidigm.com). Genes and primers are provided in Supplemental Table S1. Phenylalanine ammonia lyase (*PAL*) and cinnamyl alcohol dehydrogenase (*CAD*) primer pairs failed with the Fluidigm system and were subsequently repeated using the Bio-Rad CFX Connect Real Time System (Bio-Rad, Inc.), following the protocol in Scully et al. (2016). Briefly, 200 ng of total RNA was used for cDNA synthesis with the Transcriptor First Strand cDNA Synthesis Kit (Roche Life Science) and RT-qPCR was conducted using SsoAdvanced SYBR Green Supermix (Bio-Rad). The Fluidigm and Bio-Rad data were analyzed using the housekeeping gene, α-tubulin, for normalization and ΔCt values which were subsequently used for statistical evaluation as described below. No template and no reverse transcription controls were included to verify the absence of DNA contamination. Three biological replicates were analyzed for each genotypic class in duplicate.

Analysis of soluble and cell wall-bound phenolic compounds from stover

Soluble and wall-bound aromatic compounds were extracted from 100 mg of stover essentially as described in Palmer et al. (2008). Briefly, 1.5% acetic acid in 50% methanol was added to stover and inverted continuously for 1 h at room temperature. After a brief centrifugation at 14,000 rpm, the supernatant (containing the soluble phenolics) was removed, and centri-vaped to dryness. Soluble phenolics present in the residue were redissolved in aqueous methanol, acidified with a few drops of 6 M HCl and back extracted thrice into ethylacetate (3 × 500 µl). Combined ethylacetate fractions were dried and processed for GC–MS as described below.

Dried cell wall pellets were resuspended in 400 µl of 4.0 M NaOH and base hydrolysis was performed for 2 h at room temperature. This solution was acidified with 6 M HCl, back extracted into ethylacetate as described for soluble phenolics. Dried residues containing soluble and wall-bound phenolics were resuspended in 50 µl of pyridine derivatized with 80 µl *N*-methyl-*N*-(trimethylsilyl) trifluoroacetamide (MSTFA, Thermo Fisher) with 0.25 mg/ml toluic acid added as an internal standard. The products were analyzed using GC–MS on an Agilent G2570A integrated GC–MS system equipped with a G2913A autoinjector module, 6850 Series II GC and a 5973 Network Mass Spectrometer (Agilent, Palo Alto, CA). Relative abundances of soluble and wall-bound phenolic compounds were determined by the peak areas of major ions. Normalization was performed using the peak area for the internal standard (toluic acid). Analysis was performed in duplicates on four biological replicates per genotypic class. Genotypic classes within each event (SbF5H-16a or SbF5H-19a) were run as a batch on the GC–MS and thus treated as a random variable for statistical analysis.

Thioacidolysis

Composition of lignin subunits was determined via thioacidolysis followed by GC–MS. Stover material from mature plants was washed, derivatized and analyzed as described previously in Palmer et al. (2008). 50 µl of 1.0 mg/ml 4–4′-ethyldienesbisphenol in dichloromethane was added to each sample as an internal standard. Analysis was performed in duplicate on four biological replicates per genotype class.

Phenotypic evaluation and fiber analysis

Days to inflorescence emergence were recorded for each plant. Plant height and number of tillers were measured immediately before harvest on plants grown to full maturity. Seeds were harvested at maturity and total seed weight was measured. Total number of seeds was estimated using the mass of 100 seeds and total seed mass.

Fiber analysis was performed on ground stover using a detergent digestion protocol as described by Vogel et al. (1999). Neutral detergent fiber (NDF), acid detergent fiber (ADF) and acid detergent lignin (ADL) concentrations were measured using the ANKOM 200 fiber analyzer (ANKOM Tech Co.). Relative percentage of cell wall components were calculated using component concentrations extracted on a dry weight basis. Four biological replicates were analyzed in duplicate for the stover.

Sample fixation, histochemical staining and microscopy

Midrib and stalk tissues were fixed in an ethanol:acetic acid solution (3:1 v/v) overnight, then stored in 70% ethanol, embedded in 7% agarose and 100 µM sections were cut using a Leica VT1200s vibratome (Leica Microsystems). Sections were stained for 2 min in 0.5% potassium permanganate solution, followed by 3–4 distilled water rinses, then placed in 3.0% HCl until the deep brown color was discharged from the section then immediately followed with addition of ammonium hydroxide solution (14.8 M). Sections were imaged using an Olympus BX-51 light microscope (Olympus Co.) at × 60 (leaf midrib) and × 80 (stalk) magnification.

Bomb calorimetry

Total energy content was determined using a Parr 6400 bomb calorimeter (Parr Instrument Co.). Approximately 200 mg of dried stover combined with 600 mg of mineral oil was combusted to estimate energy value per gram of dry weight. Total energy of the stover was calculated by subtracting the total energy released from combustion of mineral oil alone from the combined mineral oil and stover sample, which was standardized to the sample weight. Analysis was performed on four biological replicates per genotypic class.

Statistical analysis

Statistical analysis of results from phenolics analysis, RT-qPCR, fiber analysis, bomb calorimetry, agronomic evaluations and thioacidolysis analysis were performed using JMP 12.2.0 (SAS Institute) mixed models. Data were tested for normality using the Wilks–Shapiro test and were log transformed if the data failed to meet normality. Pairwise comparisons among lines were performed using Tukey's Honest Significant Differences test at $\alpha \leq 0.05$.

Results

Identification and overexpression of *SbF5H*

In order to alter lignin composition and content in sorghum, the branch point controlling the conversion of G to S subunits was investigated by identifying *F5H* genes from sorghum. The sorghum genome (*S. bicolor* v3.1.1, Phytozome https://phytozome.jgi.doe.gov) contains three *F5H* genes (Sobic.001G196300.1, Sobic.002G029700.1 and Sobic.005G088400.1). Both Sobic.002G029700.1 and Sobic.005G088400.1 were not highly expressed across multiple tissue types throughout sorghum development (Supplemental Fig. 1A). However, Sobic.001G196300.1 is highly expressed across many tissue types throughout development (Supplemental Fig. 1A), which suggested that it is the major *F5H* gene involved in lignin biosynthesis. The sequences for *F5H* from five plant species (switchgrass, maize, sorghum, rice, brachypodium and Arabidopsis) were downloaded for the analysis of phylogenetic relationships. The *F5H* genes for switchgrass and Arabidopsis have been previously investigated in the effects of overexpression on lignin (Meyer et al. 1998; Wu et al. 2019). The phylogenetic analysis demonstrates the close similarity between Sobic.001G196300.1 and other monocot *F5H* genes (Supplemental Fig. 1B). Therefore, *SbF5H* (Sobic.001G196300.1) was overexpressed under the control of the constitutive E35S promoter in sorghum background RTx430. A total of seven independent transformation events were generated through *A. tumefaciens* mediated transformation, and *F5H* gene expression in these events was determined via RT-qPCR. Two independent transgenic events of *35S::SbF5H*, 16a and 19a, were selected based on gene expression level (Supplemental Fig. 2). These lines were crossed to *bmr12-ref* (RTx430) to investigate the effect of *F5H* overexpression combined with loss of *COMT* activity. Plants belonging to four genotypic classes [WT, WT(16a) and WT(19a); *35S::SbF5H*, SbF5H-16a and SbF5H-19a; *bmr12*, *bmr12*(16a) and *bmr12*(19a); *bmr12 35S::SbF5H*, stacked(16a) and stacked(19a)] from the F_2 generation for each transgenic event (16a and 19a) were further analyzed.

Expression levels of monolignol genes in sorghum plants

The expression levels of *F5H* and COMT in plants belonging to the eight different genotypes were analyzed by quantitative RT-PCR (Fig. 2). *F5H* was expressed at low basal levels in stalks of WT and *bmr12* plants and expression was significantly increased approximately 45 and 50-fold relative

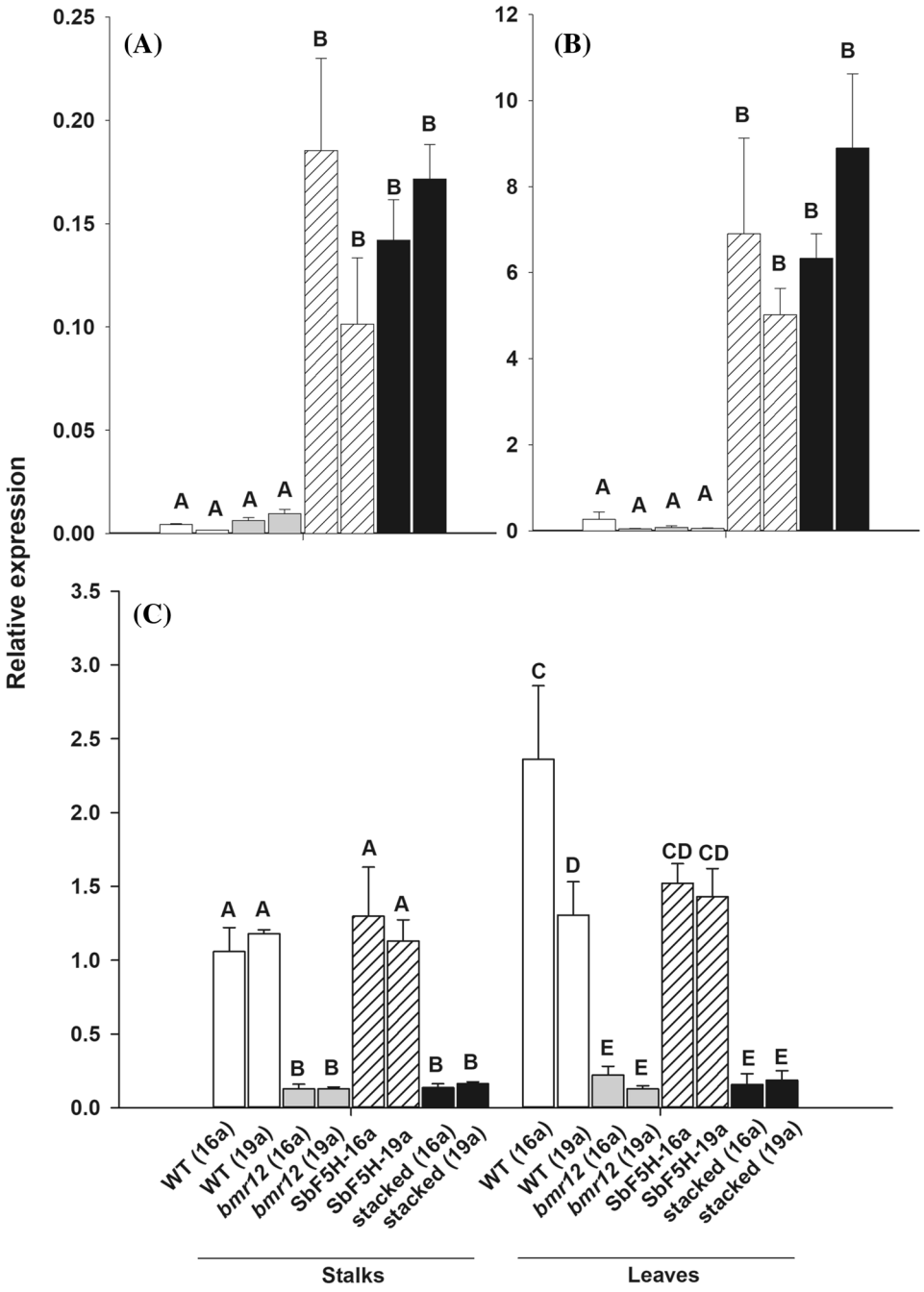


Fig. 2 Relative expression levels of (**A** and **B**) *F5H* and (**C**) *COMT* in stalk and leaves wild-type, *bmr12*, *35S::SbF5H* and stacked (*bmr12 35S::SbF5H*) plants determined using Bio-Mark HD, a microfluidic-based, high throughput qPCR system. Two independent transgenic events of *35S::SbF5H* 16a and 19a were crossed to *bmr12-ref* and four genotypic classes of F_2 generation were analyzed. Relative expression was determined using the ΔCt method with the α-tubulin gene (Sobic.001G1070200.1) for normalization. Error bars represent standard error. Samples with different letters are statistically different at $\alpha \leq 0.05$ using Tukey's HSD test

to WT for SbF5H-16a and SbF5H-19a, respectively, and increased 35 and 86-fold for stacked(16a) and stacked(19a) relative to WT, respectively ($p < 0.0001$; Fig. 2A). Similar to stalks, transcript abundance of *F5H* was low in leaves from WT and *bmr12* plants (Fig. 2B). However, *F5H* levels were elevated approximately 26 and 105-fold relative to WT in SbF5H-16a and SbF5H-19a plants, respectively and 24 and 160-fold for stacked(16a) and stacked(19a) relative to WT plants, respectively ($p < 0.0001$; Fig. 2B). Caffeic acid *O*-methyltransferase (*COMT*; *Bmr12*) expression was significantly elevated in WT and the *SbF5H* overexpression lines and decreased in stalk tissue of *bmr12*(16a) and *bmr12*(19a) with an approximately 8 and 9-fold decrease relative to WT and 8 and 11-fold decrease in stacked(16a) and stacked(19a) relative to WT, respectively ($p < 0.0001$; Fig. 2C). *COMT* expression was also significantly reduced in leaf for *bmr12*(16a) and *bmr12*(19a) with an approximately 11 and 10-fold decrease relative to WT and a 15 and 7-fold decrease in stacked(16a) and stacked(19a) relative to WT plants ($p < 0.0001$; Fig. 2C). Similar reductions in transcript abundance were previously observed in *bmr12* stalks and leaves relative to WT (Sattler et al 2012).

Transcript accumulation of additional genes associated with monolignol biosynthesis were quantified in leaves and stalks of the eight genotypes using qPCR (Fig. 3; Supplemental Table S2). Expression of 4-coumarate-CoA-ligase (*4CL*; *Bmr2*) was moderately increased in *bmr12*(16a), *bmr12*(19a), stacked(16a) and stacked(19a) compared to WT with an approximately 2.0, 3.0, 2.0 and 2.0-fold increase, respectively ($p < 0.0001$; Fig. 3). Expression of *p*-coumaroyl quinate/shikimate 3′-hydroxylase (*C3′H*) was increased approximately 1.0, 3.0, 1.0 and 3.0-fold in *bmr12*(16a), *bmr12*(19a), stacked(16a) and stacked(19a) compared to WT, respectively ($p < 0.0001$; Fig. 3). Caffeoyl CoA *O*-methyltransferase (*CCoAOMT*) transcripts were elevated in *bmr12*(16a), *bmr12*(19a), stacked(16a) and stacked(19a) compared to WT with an approximately 1.0-fold increase ($p < 0.0001$; Fig. 3). Cinnamyl CoA reductase (CCR) was increased with an approximately 1.5, 4.0, 2.0 and 4.0-fold increase in *bmr12*(16a), *bmr12*(19a), stacked(16a) and stacked(19a) relative to WT, respectively ($p < 0.0001$; Fig. 3). Likewise, the expression levels of the other monolignol genes were not significantly different in leaf tissues for these lines (Supplemental Table S2). Transcripts for a folypolyglutamate synthase (*FPGS*; SAM metabolism) and a methylenetetrahydrofolate reductase (*MTHFR*; *S*-adenosyl methionine, SAM metabolism) and 3-deoxy-o-arabino-heptulosonate phosphate synthase (*DAHP*; aromatic amino acid synthesis) were analyzed by qPCR with no significant changes detected across the genotypic classes and the tissue types tested (Supplemental Table S2).

Over-expression of *SbF5H* in the *bmr12* background affected plant growth and development

Agronomic traits were measured during the maturation of greenhouse grown plants. The leaf midribs of *SbF5H* overexpression lines were light green in color, which were similar to WT midribs (Fig. 4). However, *bmr12* leaves had reddish-brown midribs, which were previously described (Porter et al. 1978) (Fig. 4). The leaf midribs of the stacked leaves resembled the reddish-brown midribs observed on *bmr12* leaves (Fig. 4). Time to flowering for *SbF5H* overexpression lines was not significantly different from WT (Table 1). The *bmr12* and stacked plants had a significant delay in inflorescence emergence compared to both WT and *SbF5H* overexpression lines ($p = 0.0028$; Table 1). Total number of seeds and 100-seed weight were measured

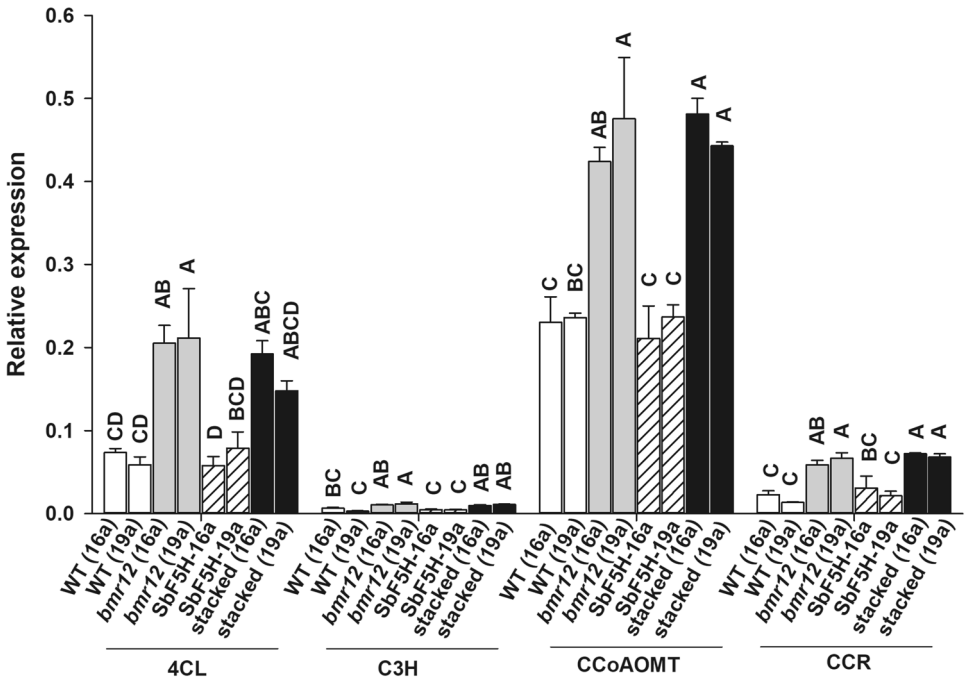


Fig. 3 Relative expression levels of other monolignol genes (*4CL* 4-coumarate-CoA ligase, *C3′H* *p*-coumaroyl quinate/shikimate 3′-hydroxylase, *CCoAOMT* caffeoyl-CoA *O*-methyltransferase, *CCR* cinnamoyl CoA reductase) that were significantly different in stalk among wild-type (WT), *bmr12*, *35S::SbF5H* and stacked (*bmr12 35S::SbF5H*) plants determined using BioMark HD, a microfluidic-based, high throughput qPCR system. Two independent transgenic events of *35S::SbF5H* 16a and 19a were crossed to *bmr12-ref* and four genotypic classes of F_2 generation were analyzed, genotypic classes are referred to as WT(16a/19a), *bmr12*(16a/19a), *SbF5H*-16a/19a and stacked(16a/19a). Relative expression was determined using the ΔCt method with the α-tubulin gene (Sobic.001G1070200.1) for normalization. Error bars represent standard error. Samples with different letters are statistically different at $\alpha \leq 0.05$ using Tukey's HSD test

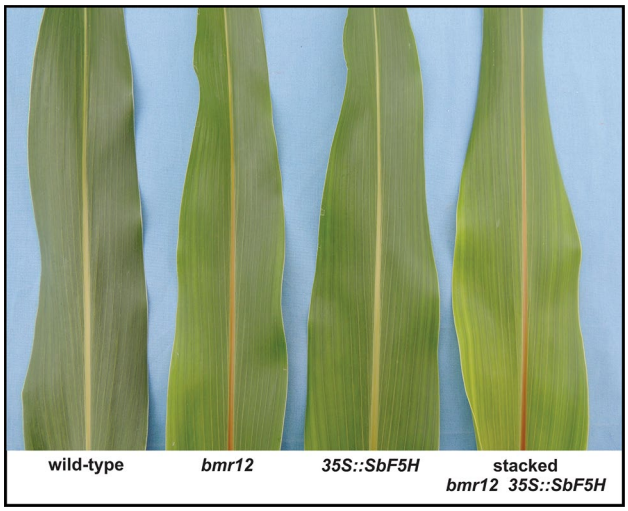


Fig. 4 The leaf midrib phenotype of the wild-type, *bmr12*, *35S::SbF5H* and stacked (*bmr12 35S::SbF5H*)

as components of grain yield. Total number of seeds estimated was significantly different across genotypic classes ($p = 0.0009$), with a significant decrease in *bmr12*(19a), but this effect was not observed for *bmr12*(16a) relative to the corresponding WT (Table 1). Grain yield was significantly decreased in both sets of stacked plants compared to WT with a 50–74% decrease (Table 1). Based on the 100-seed weight, grain filling was not significantly different in stacked plants compared to the other classes ($p = 0.8309$). We observed no significant changes in overall plant height or number of tillers produced associated with *bmr12*, overexpression of *SbF5H* or their combination in stacked plants (Table 1). Overall, combining *bmr12* with *35S::F5H* reduced grain yield through the reduction of the number seeds produced, but did not change individual seed weight relative to the value observed for WT plants.

Effects of *SbF5H* overexpression and *bmr12* on lignin content and composition

To determine whether overexpression of *SbF5H* alone or in combination with *bmr12* affected cell wall composition, fiber analysis was performed on biomass collected from *35S::SbF5H*, *bmr12*, stacked and WT plants to measure the concentrations of NDF, ADF and ADL. Levels of NDF and ADF were not significantly different across genotypic classes ($p = 0.38$ and 0.46; Fig. 5A, B). However, levels of ADL in mature stover collected from *bmr12* and stacked plants were significantly decreased by 30–48% relative to WT ($p < 0.0042$; Fig. 5C). Lignin concentration (ADL) of the stover from the *35S::SbF5H* plants was not significantly different from WT.

Table 1 Total energy and agronomic traits of wild-type (WT), *bmr12*, *35S::SbF5H* and stacked (*bmr12 35S::SbF5H*) plants

	Total energy			Agronomic traits								
	Number of individuals	Calories (g^{-1})	± SE	Inflorescence emergence (days)	± SE	Plant height (cm)	± SE	Number of tillers	± SE	100-seed weight (g)	± SE	Total number of seeds estimated	± SE
WT(16a)	4	3711.87	66.74	103^{ab}	1.84	95.75	5.56	2.5	0.65	3.36	0.19	872^{ab}	163
WT(19a)	4	3799.73	18.14	100^{b}	1.18	107	5.21	0.75	0.48	3.44	0.15	1075^{a}	79
bmr12(16a)	4	3681.73	9.42	109^{ab}	1.65	107.5	1.04	0.75	0.25	3.35	0.17	833^{ab}	117
bmr12(19a)	4	3741.56	26.35	112^{a}	4.03	100	4.26	1.5	0.29	3.17	0.35	499^{bc}	130
SbF5H-16a	4	3755.95	17.71	100^{b}	0.87	93.75	3.75	1.25	0.25	3.5	0.15	822^{ab}	149
SbF5H-19a	4	3775.51	27.92	101^{ab}	2.17	95	5.2	1.25	0.63	3.54	0.21	689^{abc}	53
Stacked(16a)	4	3741.39	25.04	112^{a}	3.87	89.75	5.27	1.5	0.48	3.24	0.19	276^{c}	108
Stacked(19a)	4	3740.91	31.84	107^{ab}	2.14	95.5	2.02	1.25	0.25	3.58	0.05	434^{bc}	81
p-value		0.31		0.0028		0.0775		0.1423		0.8309		0.0009	

Two independent transgenic events of *35S::SbF5H* 16a and 19a were crossed to *bmr12-ref* and four genotypic classes of F_2 generation were analyzed, genotypic classes are referred to as WT(16a/19a), *bmr12*(16a/19a), *SbF5H*-16a/19a and stacked(16a/19a)

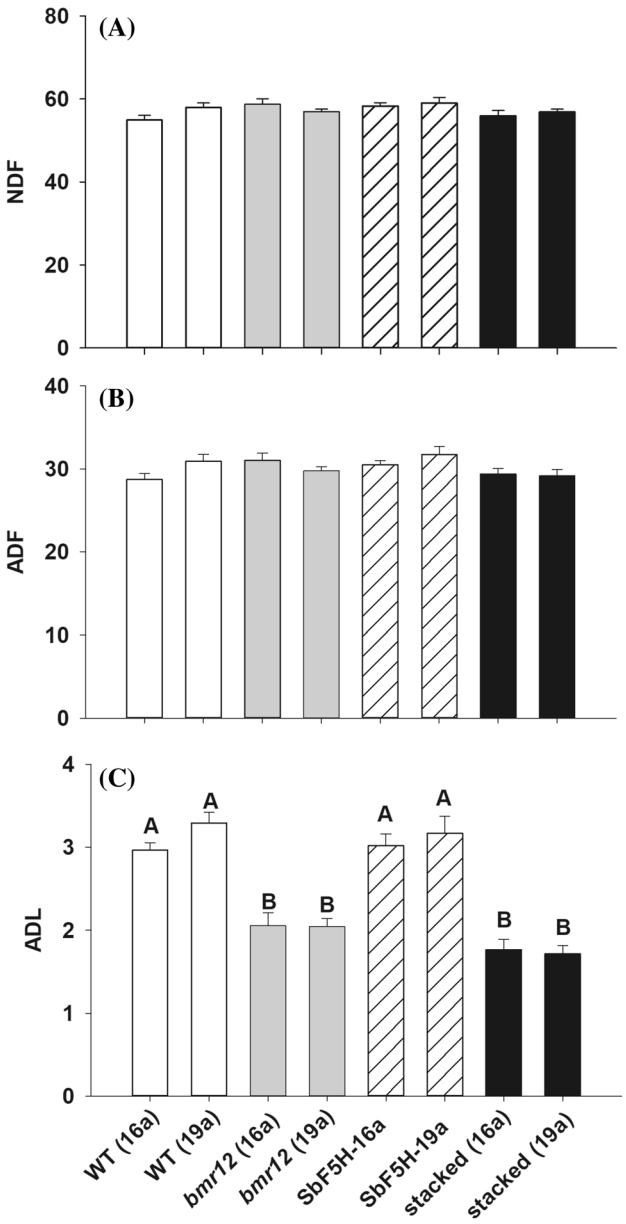


Fig. 5 Fiber analysis of mature stover from wild-type (WT), wild-type, *bmr12*, *35S::SbF5H* and stacked (*bmr12 35S::SbF5H*) for (**A**) neutral detergent fiber (NDF), (**B**) acid detergent fiber (ADF) and (**C**) acid detergent lignin (ADL). NDF, ADF and ADL determined using ANKOM fiber analyzer. Two independent transgenic events of *35S::SbF5H* 16a and 19a were crossed to *bmr12-ref* and four genotypic classes of F_2 generation were analyzed, genotypic classes are referred to as WT(16a/19a), *bmr12*(16a/19a), *SbF5H*-16a/19a and stacked(16a/19a).Values presented are least square means; error bars represent standard error. Samples with different letters are statistically different at $\alpha \leq 0.05$ using Tukey's HSD test

To evaluate changes to lignin composition, thioacidolysis was used to measure the relative amounts of β-O-4 linked monomers in the lignin (Fig. 6). The levels of H-lignin were significantly different among the genotypic classes ($p = 0.013$), however this difference was not consistent for any particular genotypic class; H-lignin was increased 1-fold in SbF5H-19a stover and decreased 1-fold in stacked(16a) stover compared to other classes. Levels of 5-OH-G-lignin, S-lignin and the ratio of S/G lignin were significantly different among genotypic classes ($p < 0.0001$). Levels of 5-OH-G-lignin were significantly elevated in stover collected from both *bmr12* and stacked plants relative to WT, ranging from a 10 to 30-fold increase (Fig. 6). The levels of G-lignin were significantly different ($p = 0.042$); similar to H-lignin, G-lignin differences were not consistent among genotypic classes, instead the change observed was higher in WT(19a), *bmr12*(16a), *bmr12*(19a) and SbF5H-19a compared to the stacked(16a) with approximately 0.6-fold increase (Fig. 6). S-lignin levels were significantly decreased in *bmr12* and stacked plants compared to WT (Fig. 6). Specifically, S-lignin ranged from a 0.6 to 0.9-fold decrease in *bmr12* and stacked stover compared to WT. S-lignin in the overexpression line, SbF5H-19a was increased 2.0 and 3.0-fold compared to WT(19a) and WT(16a), respectively, consistent with the elevated *F5H* expression levels in these plants (Fig. 2; Supplemental Fig. 3). The significant decrease in S-lignin observed in the *bmr12* and stacked plants were associated with a 0.6 to 0.8-fold decrease in the S/G ratio compared with WT. The *SbF5H* overexpression lines had a 0.9 to 1.1-fold increase in S/G ratio compared to WT. Thus, increased expression of *35S::SbF5H* in sorghum led to significant changes in the amount and ratio of lignin monomers.

To assess how these changes in lignin composition altered secondary cell walls, cross-sections of stalks and leaf midribs were observed using Maüle staining (Crocker 1933; Towers and Gibbs 1953). Cell walls with higher S-lignin content stain different shades of red. In the stalk tissue, each of the four genotypic classes was distinguishable with Maüle staining (Fig. 7). Maüle staining showed a change from light pink in WT leaf midribs to burgundy red around vascular bundles in the *SbF5H* overexpression lines (Fig. 7), which is indicative of an increase in S-lignin. In stalk, vascular bundles in *SbF5H* overexpression lines did not appear as red with the Maüle stain as in the leaf midrib, which was consistent with greater *F5H* expression in leaf tissue than in stalk (Fig. 2). The leaf midrib and stalk sections of *bmr12* stained a yellow–brown color around vascular bundles, which may indicate a depletion in S-lignin subunits (Fig. 7). The leaf midrib staining pattern of the stacked plants was similar to *bmr12* pattern, which also may indicate the depletion of S-lignin, compared to the yellow–brown coloration of WT leaf midrib stained sections. However, vascular bundles from the stacked stalks stained a distinct burgundy-red to dark-brown color. Although this staining was different in appearance than the other classes, similar colorations have been previously associated with low levels of S-lignin (Crocker 1933; Sewalt et al. 1997).

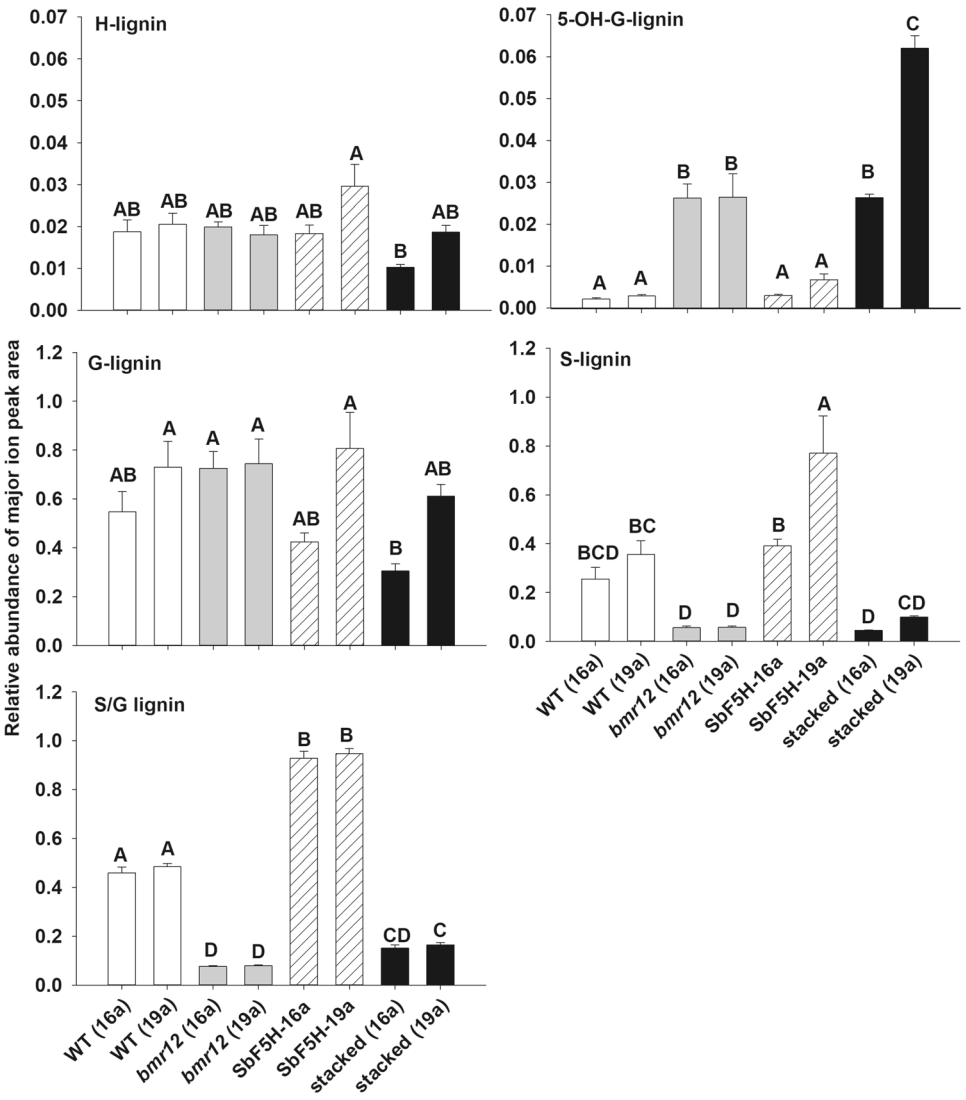


Fig. 6 Lignin composition determined using thioacidolysis for wild-type (WT), wild-type, *bmr12*, *35S::SbF5H* and stacked (*bmr12 35S::SbF5H*) stover. The abundance of major ion peak area of *p*-hydroxyphenyl (H-lignin), 5-hydroxy-guaiacyl (5-OH-G-lignin), guaiacyl (G-lignin) and syringyl (S-lignin) subunits relative to the internal standard (4,4′-ethylidenebisphenol) were determined by GC/MS. Two independent transgenic events of *35S::SbF5H* 16a and 19a were crossed to *bmr12-ref* and four genotypic classes of F_2 generation were analyzed, genotypic classes are referred to as WT(16a/19a), *bmr12*(16a/19a), *SbF5H*-16a/19a and stacked(16a/19a). Values presented represent least square means; error bars represent standard error. Samples with different letters are statistically different at $\alpha \leq 0.05$ using Tukey's HSD test

To determine how overexpression of *35S::SbF5H* and *bmr12* affected the cell wall phenylpropanoid constituents separate from lignin, soluble and cell wall-bound phenolic compounds were extracted from mature stover and analyzed by GC–MS. Significantly higher levels of syringic acid and sinapic acid in the soluble phenolic compounds were detected in the *SbF5H* overexpression and stacked lines relative to WT ($p < 0.0001$; Fig. 8). Levels of ferulic, 4-coumaric and vanillic acid in soluble phenolic fraction were significantly decreased in the *SbF5H* overexpression and stacked lines relative to WT ($p = 0.0004, < 0.0001$ and < 0.0001, respectively; Fig. 8). Soluble 5-hydroxy ferulic acid (5-OH ferulic acid) was significantly increased in the *bmr12* and stacked plants ($p = 0.0065$; Fig. 8). Caffeic acid was not significantly different in the soluble phenolics among genotypic classes ($p = 0.5624$; Fig. 8). Relative abundances of cell wall-bound phenolic esters are shown in Fig. 9. Sinapic acid and syringic acid levels were increased in plants overexpressing *SbF5H* relative to WT ($p < 0.0001$; Fig. 9). While levels of syringic acid were significantly decreased in *bmr12*, they were not significantly altered in the stacked stover relative to WT ($p < 0.0001$; Fig. 9). Levels of 4-coumaric acid were significantly decreased in *bmr12* and stacked relative to WT and *SbF5H* overexpression lines ($p < 0.0001$; Fig. 9). Cell wall-bound 5-OH ferulic acid was significantly greater in all *bmr12* plants ($p < 0.0001$; Fig. 9). Wall-bound ferulic acid was slightly increased in *bmr12* and stacked stover, but changes in caffeic acid and vanillic acid were inconsistent among genotypic classes (Fig. 9). Similarly, consistent differences between the genotypic classes for other phenolic compounds identified by GC–MS in soluble or wall-bound stover extracts were not observed (Supplemental Table S3). Overexpression of *SbF5H* increased levels of wall-bound and soluble sinapic acid even in the *bmr12* background, unlike its lack of effect on S-lignin levels in *bmr12* plants.

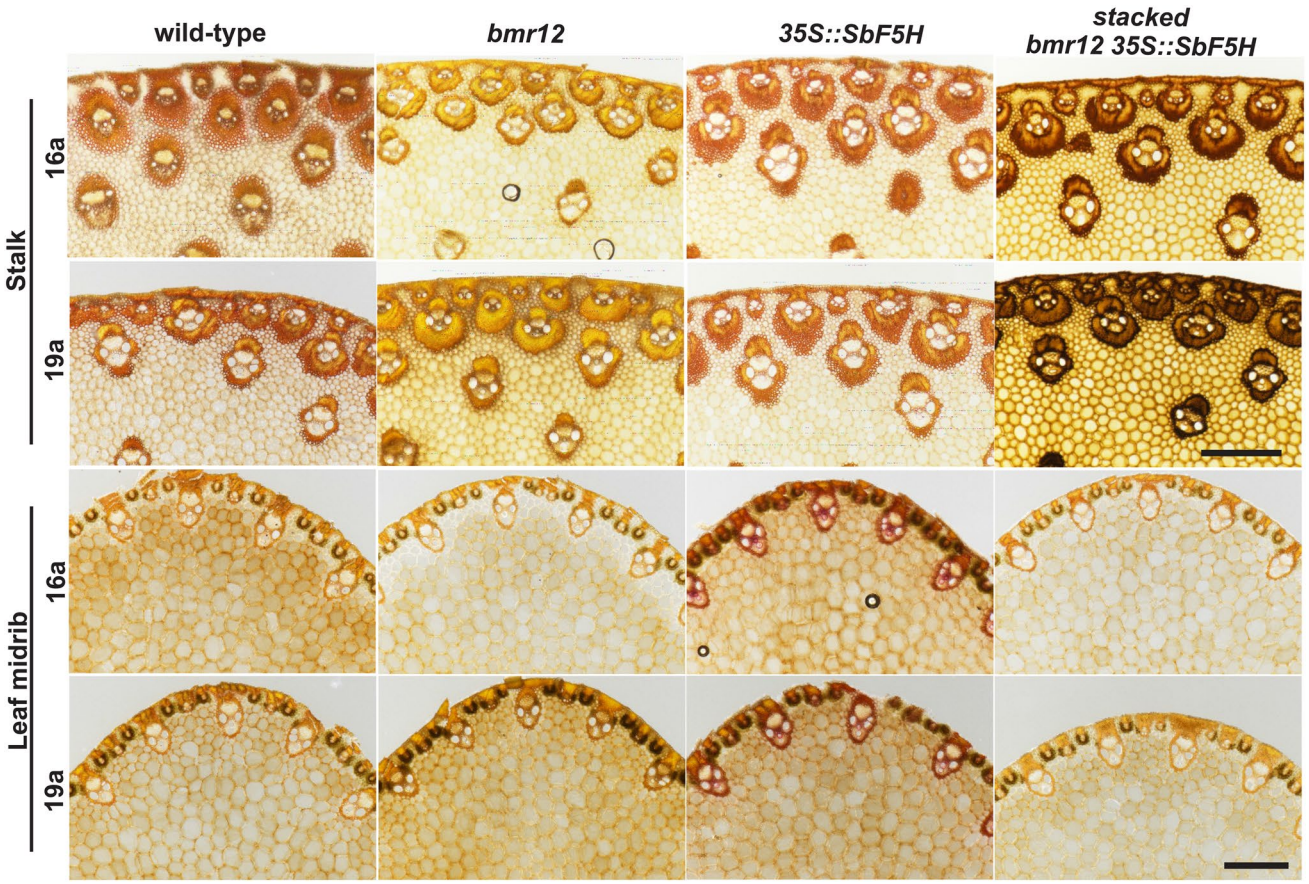


Fig. 7 Visualization of syringyl (S)-lignin subunits after Maüle staining of stalk and leaf midrib tissue taken from wild-type, *bmr12*, *35S::SbF5H* and stacked (*bmr12 35S::SbF5H*) plants. Two independent transgenic events of *35S::SbF5H* 16a and 19a were crossed to *bmr12-ref* and four genotypic classes of F_2 generation were analyzed, genotypic classes are referred to as WT(16a/19a), *bmr12*(16a/19a), *SbF5H*-16a/19a and stacked(16a/19a). Scale bar = 500 µm for both stalk and leaf midrib. Stalk tissue was observed at × 80 magnification, leaf midrib was observed at × 60 magnification using an Olympus BX-51 light microscope (Olympus Co.)

Discussion

The effects of *SbF5H* overexpression (Sobic.001G196300.1) alone and in combination with *bmr12* were assessed on agronomic and cell wall characteristics in sorghum plants. F5H is a cytochrome P450 hydroxylase that catalyzes the conversion of coniferylaldehyde to 5-hydroxyconiferylaldehyde (Humphreys et al. 1999), a critical intermediate required for the eventual biosynthesis of sinapyl alcohol. Sinapyl alcohol is the precursor for the formation of S-lignin, a more linear form of lignin polymer compared to G-lignin (Guo et al. 2001). Sorghum has three copies of genes encoding putative F5H, of these only Sobic.001G196300.1 was expressed with certainty in lignifying tissues and therefore chosen for overexpression. COMT encoded by *Bmr12* is the enzyme that utilizes the product of F5H catalysis to form sinapylaldehyde. Sorghum *bmr12* mutants are devoid of a functional COMT and deficient in S-lignin (Bout and Vermerris 2003; Sattler et al. 2012; Eudes et al. 2017). Two independent *SbF5H* overexpression lines were crossed with *bmr*12 to produce *bmr12 35S::SbF5H* (stacked) plants to ascertain the consequences of increased availability of the COMT substrate on biomass composition. The results of this study demonstrated that overexpression of *SbF5H* increased S-lignin monomers and the ratio of S-lignin to G-lignin monomers as well as stimulated biosynthesis of sinapic and syringic acid in both soluble and cell wall-bound fractions from sorghum stover. The results of this study further demonstrated that overexpression of *SbF5H* does not compensate for loss of COMT activity in *bmr12* in regard to lignin concentration and composition; both S-lignin and lignin (ADL) concentrations were unchanged in stacked plants compared to *bmr12*.

Lignin biosynthesis is plastic, and both its content and monomeric composition have been substantially manipulated through genetic and transgenic approaches in plants (Zhao et al. 2002; Reddy et al. 2005; Weng et al. 2010; Takeda et al. 2017; Wu et al. 2019). Alterations of syringyl

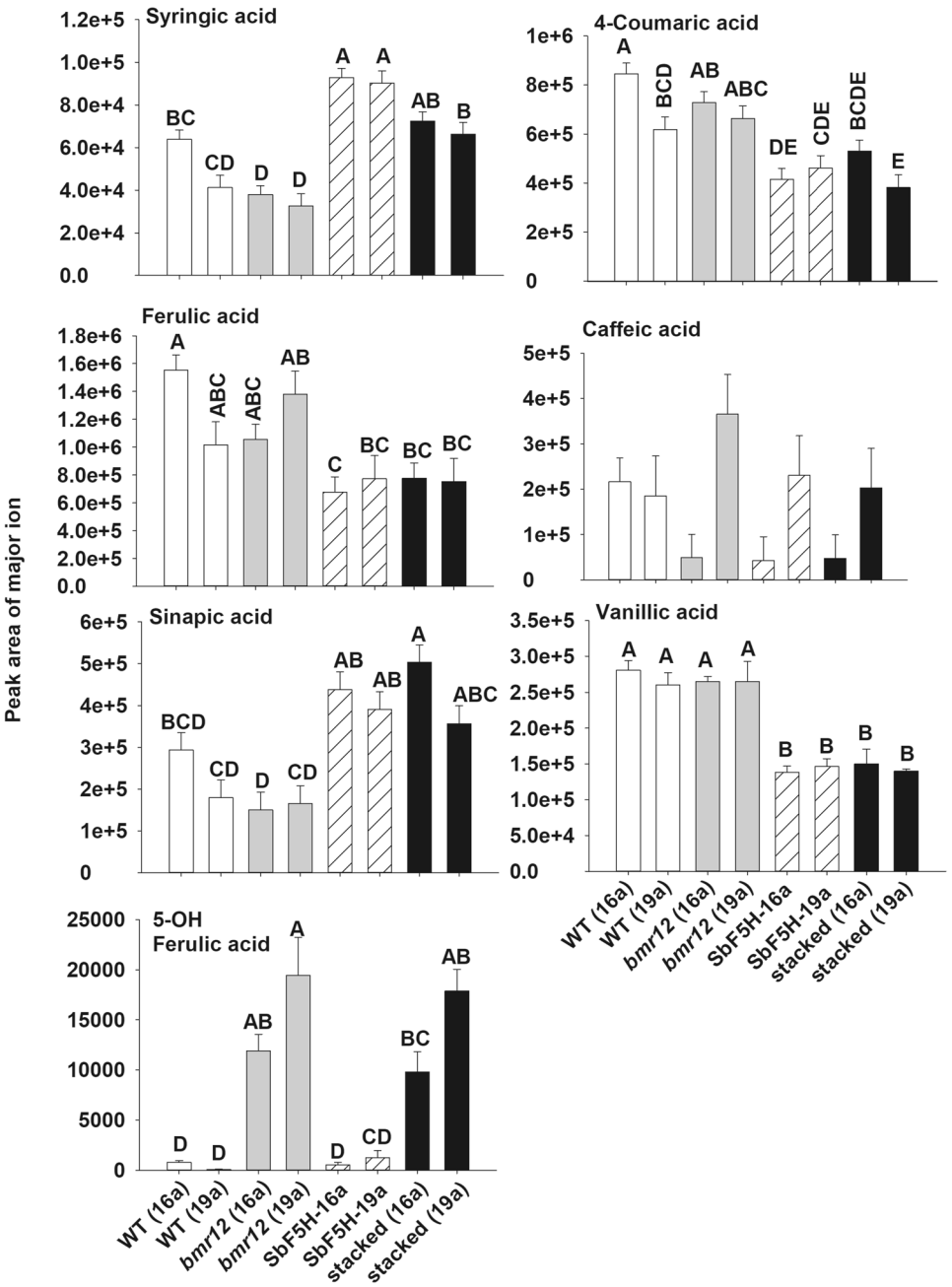


Fig. 8 Abundance of soluble phenolic acids in stover from wild-type (WT), *bmr12*, *35S::SbF5H* and stacked (*bmr12 35S::SbF5H*) plants. Two independent transgenic events of *35S::SbF5H* 16a and 19a were crossed to *bmr12-ref* and four genotypic classes of F_2 generation were analyzed, genotypic classes are referred to as WT(16a/19a), *bmr12*(16a/19a), *SbF5H*-16a/19a and stacked(16a/19a). Phenolic acids were analyzed via GC–MS. Values presented are least square means; error bars represent standard error. Samples with different letters for wall-bound and soluble fractions are statistically different from one another at $\alpha \leq 0.05$ using Tukey's HSD test

lignin concentration have been achieved through the manipulation of F5H and COMT, which are two enzymes required to convert coniferoyl groups to sinapoyl groups. For example, *FAH1* encodes the Arabidopsis *F5H* gene, and the *fah1* mutant had a major decrease in S-lignin and an enrichment in G-lignin (Chapple et al. 1992; Meyer et al. 1998). Alternatively, overexpression of *F5H* increased S-lignin and decreased G-lignin in *Arabidopsis*, tobacco, switchgrass and rice (Franke et al. 2000; Weng et al. 2010; Takeda et al. 2017; Wu et al. 2019). Consistent with previous studies, overexpression of *SbF5H* significantly altered S-lignin levels, which resulted in a significant increase of S/G ratio and indicated that overexpression of *SbF5H* is sufficient to increase the flux toward S-lignin synthesis in sorghum. Phenotypic effects of increasing S-lignin levels were largely correlated with expression levels of *SbF5H*, which varied among the two transgenic events included in this study. Elevated levels of S-lignin were observed in SbF5H-19a, which had the higher level of *F5H* expression, specifically in leaf tissue. Maüle staining is commonly used to determine differences in lignin composition because it preferentially stains S-lignin monomers red in color (Towers and Gibbs 1953; Pradhan Mitra and Loqué 2014). Therefore, the increased red coloration observed around the vascular bundles in leaf

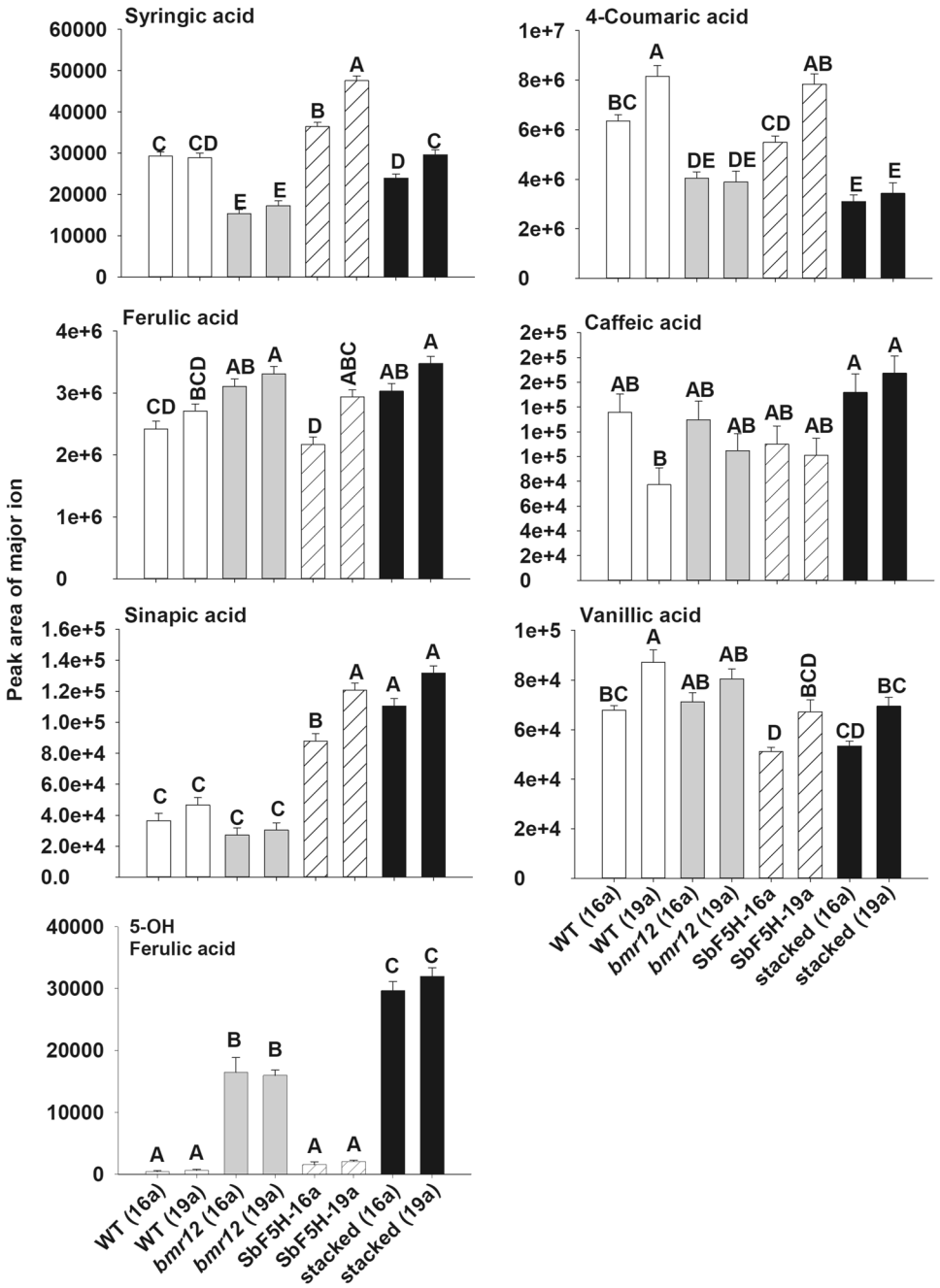


Fig. 9 Abundance of wall-bound phenolic acids in stover from wild-type (WT), *bmr12*, *35S::SbF5H* and stacked (*bmr12 35S::SbF5H*) plants. Two independent transgenic events of *35S::SbF5H* 16a and 19a were crossed to *bmr12-ref* and four genotypic classes of F_2 generation were analyzed, genotypic classes are referred to as WT(16a/19a), *bmr12*(16a/19a), *SbF5H*-16a/19a and stacked(16a/19a). Phenolic acids were analyzed via GC–MS. Values presented are least square means; error bars represent standard error. Samples with different letters for wall-bound and soluble fractions are statistically different from one another at $\alpha \leq 0.05$ using Tukey's HSD test

midrib in *SbF5H* overexpression lines visually demonstrated the increased deposition of S-lignin. Coniferyl alcohol is monomethoxylated while sinapyl alcohol is dimethoxylated and this difference causes G-lignin to be more branched than the relatively linear S-lignin (Li et al. 2016). Hence, the amount of G-lignin is often negatively correlated with in vitro digestibility of forage crops, which is consistent with the more condensed structure of G-lignin (Jung and Deetz 1993). When *F5H* was overexpressed in poplar, S-lignin subunits and the S/G ratio significantly increased and subsequently greatly increased pulping efficiencies (Huntley et al. 2003). Here, we have similarly engineered sorghum with increased S-lignin subunits and S/G ratio by overexpression of *SbF5H*. In addition, the overexpression of *SbF5H* in sorghum did not significantly affect growth and development. However, when overexpression of *SbF5H* was combined with *bmr12* there was a significant decrease in seed production. In Arabidopsis, overexpression of *F5H* in a COMT-deficient background affected pollen formation, which resulted in male-sterility (Weng et al. 2010). Overexpression of *SbF5H* in *bmr12* also appeared to affect fertility, but grain development was unaffected based on comparisons

of seed weight and size with WT. In addition, flowering was delayed in both the *bmr12* and stacked (*bmr12 35S::SbF5H*) plants, which was observed previously in *bmr12-ref*, but not in other alleles of *Bmr12* (Vermerris et al. 2002; Bout and Vermerris 2003; Sattler et al. 2010).

As previously reported, S-lignin was significantly decreased in *bmr12* in the current study and S-lignin levels were also decreased in the stacked plants, which indicated that overexpression of *SbF5H* via 35S promoter is unable to circumvent the loss of COMT activity (*bmr12-ref*) in the biosynthesis of S-lignin monolignols. In addition to altering lignin composition, total lignin concentrations were also decreased in *bmr12* and stacked stover. Loss of COMT function also had resulted in the deposition of unique 5-OH-G-lignin due to incorporation of 5-hydroxyconiferyl alcohol as previously observed in Arabidopsis, maize, sorghum, and tobacco (Piquemal et al. 2002; Zhao et al. 2002; Palmer et al. 2008; Weng et al. 2010). This decreased deposition of S-lignin within the cell walls of the *bmr12* and stacked in leaf midribs and stalks was visible and distinguishable from the other classes of plants using Maüle staining in the current study. However, the *bmr12* vascular bundles stained distinctly yellow, while stacked vascular bundles stained dark brown, which indicates there are detectable differences in the cell wall composition of *bmr12* and *bmr12* with overexpression of *SbF5H*. Although thioacidolysis did not detect obvious differences in the amounts of S-lignin within the cell walls, elevated levels of 5-hydroxy-coniferyl groups in the stacked plants relative to *bmr12* may explain the different staining patterns. The results from thioacidolysis and Maüle staining indicate that the combination of *bmr12* and overexpression of *SbF5H* alters phenylpropanoid composition of the cell walls, which does not involve the thioacidolysis-detected β-O_4 linked S-lignin residues (Yue et al. 2012, 2017). In addition, this study shows that Maüle staining of sorghum tissues is a sensitive tool to identify changes to phenylpropanoids within cell walls.

Overexpression of *SbF5H* alone did not change the lignin concentration, but it altered the accumulation of soluble and wall-bound phenolic compounds in the stover. Both *SbF5H* transgenic lines accumulated lower levels of ferulic acid and vanillic acid in the soluble fraction of stover, which are derived from substrates of F5H. In addition, there was increased accumulation of soluble sinapic acid in the stover of *SbF5H* overexpression lines, which supports the hypothesis that F5H may utilize ferulic acid as a substrate leading to the synthesis of sinapic acid (Meyer et al. 1998). In *bmr12*, the levels of 5-hydroxyferulic acid in both soluble and wall-bound stover fractions were increased relative to WT, and these levels were further augmented in the stacked stover. Similar to sorghum, Arabidopsis *comt* mutants uniquely had 5-hydroxyferulic acid formation when compared to WT and other mutants of genes in monolignol biosynthesis (Vanholme et al. 2012). In contrast, the levels of wall-bound syringic and 4-coumaric acid were decreased in *bmr12* plants in the current study. Similarly, 4-coumaric and syringic acid were decreased in *bmr12* (Tx430) relative to WT in stalk internode tissue (Palmer et al. 2008). Previously, sinapic acid was shown to be increased in the soluble fraction from sorghum internodes (Palmer et al. 2008). Although in the current study there was no observation of an increase in soluble sinapic acid levels in *bmr12*, both studies observed no major reduction in wall-bound sinapic acid, unlike the reductions of S-lignin. Taken together, these data indicated that sinapic acid synthesis does not require the activity of Bmr12 (COMT), in contrast to S-lignin synthesis.

Similar to the overexpression of *SbCCoAOMT* previously observed in sorghum (Tetreault et al. 2018), overexpression of the single pathway gene *SbF5H* did not affect the expression of other genes related to monolignol biosynthesis. Leaf and stalk tissues from *bmr12* and stacked had a substantial decrease in *COMT* expression, as expected (Bout and Vermerris 2003; Sattler et al. 2012). However, *bmr12* impacted the expression of four genes encoding enzymes upstream in the monolignol biosynthesis pathway, whose expression were mainly increased relative to WT or *SbF5H* overexpression in stalks. In *bmr12*, the consequence of decreased COMT activity appeared to increase expression of genes encoding the upstream enzymes in monolignol biosynthesis to compensate for the reduced flux to sinapyl alcohol. The expression of genes encoding enzymes involved in SAM metabolism (FPGS) and amino acid synthesis (DAHP) were unaffected in *bmr12*, which suggests that changes in the monolignol biosynthesis pathway in these plants may not sufficiently alter aromatic amino acid or SAM pools to warrant a transcriptional response.

In the current study cauliflower mosaic virus 35S promoter was chosen due its known expression in sorghum (Hill-Ambroz and Weeks 2001), and its ability to confer high expression levels without tissue-specificity (Odell et al. 1985). However, engineering compositional changes to cell walls for bioenergy feedstock improvement, a stalk specific promoter would allow a targeted approach to redesign lignin biosynthesis once stalk specific promoter become available for grasses. In Arabidopsis and tobacco, a C4H promoter controlling F5H resulted in significantly decreased G-lignin and increased S-lignin levels, but a decrease to G-lignin levels were not observed using 35S promoter in these plants (Meyer et al. 1998; Franke et al. 2000). If such a promoter were available for grasses, then similar results may be achieved in sorghum.

In conclusion, the present study identified the sorghum *F5H* gene and showed it can be used to modify lignin composition through transgenic strategies. In sorghum, F5H activity appeared to be the limiting step in the synthesis of S-lignin monomers, because its product 5-hydroxyconiferoyl

groups did not accumulate in the *SbF5H* overexpression lines. In addition, 5-hydroxyconiferyl acid and 5-hydroxyguaiacyl lignin only accumulated within cell walls of *bmr12* mutants and the stacked lines, which indicate that there is an overabundance of Bmr12 (COMT) activity in sorghum. This study of *SbF5H* demonstrates its utility toward designing grass biomass with greatly altered S/G ratios for bioenergy uses.

Acknowledgements We thank John Toy, Coehn Preble, Samantha Timmons and Sara Finnegan for technical assistance, and Dr. Heather Van Buskirk for critically reviewing the manuscript. We thank the Plant Transformation Core Research Facility at the University of Nebraska for creating the overexpression lines. This research was supported by the United States Department of Agriculture: National Institute of Food and Agriculture AFRI Grant Number 2011-67009-30026 (SES and DLF-H) and additional funding from USDA-ARS, CRIS Projects 3042-21220-033-00-D (SES and DLF-H). The University of Nebraska DNA Sequencing Core receives partial support from the National Institute for General Medical Science (NIGMS) INBRE-P20GM103427-14 and COBRE-1P30GM110768-01 Grants as well as The Fred and Pamela Buffett Cancer Center Support Grant: P30CA036727. This publication's contents are the sole responsibility of the authors and do not necessarily represent the official views of the NIH or NIGMS. The US Department of Agriculture, Agricultural Research Service, is an equal opportunity/affirmative action employer and all agency services are available without discrimination. Mention of commercial products and organizations in this manuscript is solely to provide specific information. It does not constitute endorsement by USDA-ARS over other products and organizations not mentioned.

Author contributions HMT, SES and DLFH designed the research; HMT, TG, NAP, SS and ZG performed the experiments; HMT, TG, NAP, DLFH, GS and SES analyzed and interpreted the data; HMT and SES wrote the first draft of the manuscript, and all authors reviewed and revised the manuscript prior to publication.

References

Boerjan W, Ralph J, Baucher M (2003) Lignin biosynthesis. Annu Rev Plant Biol 54:519–546

Bout S, Vermerris W (2003) A candidate-gene approach to clone the sorghum Brown midrib gene encoding caffeic acid *O*-methyltransferase. Mol Genet Genomics 269:205–214

Chapple C, Vogt T, Ellis BE, Somerville CR (1992) An Arabidopsis mutant defective in the general phenylpropanoid pathway. Plant Cell 4(11):1413–1424

Chen F, Dixon RA (2007) Lignin modification improves fermentable sugar yields for biofuel production. Nat Biotechnol 25:759–761

Cherney J, Cherney D, Akin D, Axtell J (1991) Potential of brown-midrib, low-lignin mutants for improving forage quality. In: Advances in agronomy. Elsevier, Amsterdam, pp 157–198

Crocker EC (1933) Maule lignin test on podocarpus wood. Bot Gaz 95:168–171

Dien BS, Sarath G, Pedersen JF, Sattler SE, Chen H, Funnell-Harris DL, Nichols NN, Cotta MA (2009) Improved sugar conversion and ethanol yield for forage sorghum (*Sorghum bicolor* L. Moench) lines with reduced lignin contents. Bioenergy Res 2:153–164

Eudes A, Dutta T, Deng K, Jacquet N, Sinha A, Benites VT, Baidoo EEK, Richel A, Sattler SE, Northen TR, Singh S, Simmons BA, Loque D (2017) SbCOMT (Bmr12) is involved in the biosynthesis of tricin-lignin in sorghum. PLoS ONE 12:e0178160

Franke R, Mcmichael CM, Meyer K, Shirley AM, Cusumano JC, Chapple C (2000) Modified lignin in tobacco and poplar plants over-expressing the Arabidopsis gene encoding ferulate 5-hydroxylase. Plant J 22:223–234

Godin B, Nagle N, Sattler S, Agneessens R, Delcarte J, Wolfrum E (2016) Improved sugar yields from biomass sorghum feedstocks: comparing low-lignin mutants and pretreatment chemistries. Biotechnol Biofuels 9:251

Guo D, Chen F, Wheeler J, Winder J, Selman S, Peterson M, Dixon RA (2001) Improvement of in-rumen digestibility of alfalfa forage by genetic manipulation of lignin *O*-methyltransferases. Transgenic Res 10:457–464

Hatfield RD, Rancour DM, Marita JM (2017) Grass cell walls: a story of cross-linking. Front Plant Sci 7:2056

Hill-Ambroz KL, Weeks JT (2001) Comparison of constitutive promoters for sorghum [*Sorghum bicolor* (L.) Moench] transformation. Cereal Res Commun 29:17–24

Humphreys JM, Hemm MR, Chapple C (1999) New routes for lignin biosynthesis defined by biochemical characterization of recombinant ferulate 5-hydroxylase, a multifunctional cytochrome P450-dependent monooxygenase. Proc Natl Acad Sci USA 96:10045–10050

Huntley SK, Ellis D, Gilbert M, Chapple C, Mansfield SD (2003) Significant increases in pulping efficiency in C4H–F5H-transformed poplars: improved chemical savings and reduced environmental toxins. J Agric Food Chem 51:6178–6183

Jorgenson LR (1931) Brown midrib in maize and its linkage relations. J Am Soc Agron 23:549–557

Jouanin L, Goujon T, De Nadaï V, Martin M-T, Mila I, Vallet C, Pollet B, Yoshinaga A, Chabbert B, Petit-Conil M, Lapierre C (2000) Lignification in transgenic poplars with extremely reduced caffeic acid *O*-methyltransferase activity. Plant Physiol 123:1363–1374

Jung HG, Deetz DA (1993) Cell wall lignification and degradability. In: Jung HG, Buxton DR, Hatfield RD, Ralph J (eds) Forage cell wall structure and digestibility. American Society of Agronomy Crop Science Society of America, Soil Science Society of America, Madison, pp 315–346

Lee DK, Aberle E, Anderson EK, Anderson W, Baldwin BS, Baltensperger D, Barrett M, Blumenthal J, Bonos S, Bouton J, Bransby DI, Brummer C, Burks PS, Chen C, Daly C, Egenolf J, Farris RL, Fike JH, Gaussoin R, Gill JR, Gravois K, Halbleib MD, Hale A, Hanna W, Harmoney K, Heaton EA, Heiniger RW, Hoffman L, Hong CO, Kakani G, Kallenbach R, Macoon B, Medley JC, Missaoui A, Mitchell R, Moore KJ, Morrison JI, Odvody GN, Richwine JD, Ogoshi R, Parrish JR, Quinn L, Richard E, Rooney WL, Rushing JB, Schnell R, Sousek M, Staggenborg SA, Tew T, Uehara G, Viands DR, Voigt T, Williams D, Williams L, Wilson LT, Wycislo A, Yang Y, Owens V (2018) Biomass production of herbaceous energy crops in the United States: field trial results and yield potential maps from the multiyear regional feedstock partnership. GCB Bioenergy 10:698–716

Li M, Pu Y, Ragauskas AJ (2016) Current understanding of the correlation of lignin structure with biomass recalcitrance. Front Chem 4:45

Meyer K, Shirley AM, Cusumano JC, Bell-Lelong DA, Chapple C (1998) Lignin monomer composition is determined by the expression of a cytochrome P450-dependent monooxygenase in *Arabidopsis*. Proc Natl Acad Sci USA 95:6619–6623

Mitchell R, Schmer M, Anderson W, Jin V, Balkcom K, Kiniry J, Coffin A, White P (2016) Dedicated energy crops and crop residues for bioenergy feedstocks in the central and eastern USA. Bioenergy Res 9:384–398

Odell JT, Nagy F, Chua N-H (1985) Identification of DNA sequences required for activity of the cauliflower mosaic virus 35S promoter. Nature 313:810–812

Oliver A, Pedersen J, Grant R, Klopfenstein T (2005) Comparative effects of the sorghum bmr-6 and bmr-12 genes: I. Forage sorghum yield and quality. Crop Sci 45:2234–2239

Palmer NA, Sattler SE, Saathoff AJ, Funnell D, Pedersen JF, Sarath G (2008) Genetic background impacts soluble and cell wall-bound aromatics in brown midrib mutants of sorghum. Planta 229:115

Pedersen JF, Funnell DL, Toy JJ, Oliver A, Grant R (2006) Registration of twelve grain sorghum genetic stocks near-isogenic for the brown midrib genes bmr-6 and bmr-12. Crop Sci 46:491–493

Piquemal J, Chamayou S, Nadaud I, Beckert M, Barrière Y, Mila I, Lapierre C, Rigau J, Puigdomenech P, Jauneau A, Digonnet C, Boudet A-M, Goffner D, Pichon M (2002) Down-regulation of caffeic acid *O*-methyltransferase in maize revisited using a transgenic approach. Plant Physiol 130:1675–1685

Porter KS, Anxtell JD, Lechtenberg VL, Colenbrander VF (1978) Phenotype, fiber composition, and in vitro dry matter disappearance of chemically induced brown midrib (bmr) mutants of sorghum. Crop Sci. https://doi.org/10.2135/cropsci1978.0011183X001800020002x

Pradhan Mitra P, Loqué D (2014) Histochemical staining of *Arabidopsis thaliana* secondary cell wall elements. J Vis Exp. https://doi.org/10.3791/51381

Ragauskas AJ, Beckham GT, Biddy MJ, Chandra R, Chen F, Davis MF, Davison BH, Dixon RA, Gilna P, Keller M (2014) Lignin valorization: improving lignin processing in the biorefinery. Science 344:1246843

Ralph J, Lapierre C, Marita JM, Kim H, Lu F, Hatfield RD, Ralph S, Chapple C, Franke R, Hemm MR, Van Doorsselaere J, Sederoff RR, O'Malley DM, Scott JT, Mackay JJ, Yahiaoui N, Boudet A-M, Pean M, Pilate G, Jouanin L, Boerjan W (2001) Elucidation of new structures in lignins of CAD- and COMT-deficient plants by NMR. Phytochemistry 57:993–1003

Ralph J, Lundquist K, Brunow G, Lu F, Kim H, Schatz PF, Marita JM, Hatfield RD, Ralph SA, Christensen JH (2004) Lignins: natural polymers from oxidative coupling of 4-hydroxyphenylpropanoids. Phytochem Rev 3:29–60

Rastogi S, Dwivedi UN (2006) Down-regulation of lignin biosynthesis in transgenic *Leucaena leucocephala* harboring O-methyltransferase gene. Biotechnol Prog 22:609–616

Reddy MSS, Chen F, Shadle G, Jackson L, Aljoe H, Dixon RA (2005) Targeted down-regulation of cytochrome P450 enzymes for forage quality improvement in alfalfa (*Medicago sativa* L.). Proc Natl Acad Sci USA 102:16573–16578

Rogers SO, Bendich AJ (1985) Extraction of DNA from milligram amounts of fresh, herbarium and mummified plant tissues. Plant Mol Biol 5:69–76

Rooney WL (2004) Sorghum improvement—integrating traditional and new technology to produce improved genotypes. Adv Agron 83:37–109

Saballos A, Ejeta G, Sanchez E, Kang CH, Vermerris W (2009) A genomewide analysis of the cinnamyl alcohol dehydrogenase family in sorghum [*Sorghum bicolor* (L.) Moench] identifies SbCAD2 as the brown midrib6 gene. Genetics 181:783–795

Saballos A, Sattler SE, Sanchez E, Foster TP, Xin Z, Kang C, Pedersen JF, Vermerris W (2012) Brown midrib2 (Bmr2) encodes the major 4-coumarate:coenzyme A ligase involved in lignin biosynthesis in sorghum (*Sorghum bicolor* (L.) Moench). Plant J 70:818–830

Sarath G, Mitchell RB, Sattler SE, Funnell D, Pedersen JF, Graybosch RA, Vogel KP (2008) Opportunities and roadblocks in utilizing forages and small grains for liquid fuels. J Ind Microbiol Biotechnol 35:343–354

Sattler SE, Saathoff AJ, Haas EJ, Palmer NA, Funnell-Harris DL, Sarath G, Pedersen JF (2009) A nonsense mutation in a cinnamyl alcohol dehydrogenase gene is responsible for the sorghum brown midrib6 phenotype. Plant Physiol 150:584–595

Sattler SE, Funnell-Harris DL, Pedersen JF (2010) Efficacy of singular and stacked brown midrib 6 and 12 in the modification of lignocellulose and grain chemistry. J Agric Food Chem 58:3611–3616

Sattler SE, Palmer NA, Saballos A, Greene AM, Xin Z, Sarath G, Vermerris W, Pedersen JF (2012) Identification and characterization of four missense mutations in brown midrib 12 (Bmr12), the caffeic O-methyltransferase (COMT) of sorghum. Bioenergy Res 5:855–865

Scully ED, Gries T, Sarath G, Palmer NA, Baird L, Serapiglia MJ, Dien BS, Boateng AA, Ge Z, Funnell-Harris DL, Twigg P, Clemente TE, Sattler SE (2016) Overexpression of SbMyb60 impacts phenylpropanoid biosynthesis and alters secondary cell wall composition in *Sorghum bicolor*. Plant J 85:378–395

Sewalt VJ, Ni W, Blount JW, Jung HG, Masoud SA, Howles PA, Lamb C, Dixon RA (1997) Reduced lignin content and altered lignin composition in transgenic tobacco down-regulated in expression of L-phenylalanine ammonia-lyase or cinnamate 4-hydroxylase. Plant Physiol 115:41–50

Stewart JJ, Akiyama T, Chapple C, Ralph J, Mansfield SD (2009) The effects on lignin structure of overexpression of ferulate 5-hydroxylase in hybrid poplar. Plant Physiol 150:621–635

Takeda Y, Koshiba T, Tobimatsu Y, Suzuki S, Murakami S, Yamamura M, Rahman MM, Takano T, Hattori T, Sakamoto M (2017) Regulation of CONIFERALDEHYDE 5-HYDROXYLASE expression to modulate cell wall lignin structure in rice. Planta 246:337–349

Tetreault HM, Scully ED, Gries T, Palmer NA, Funnell-Harris DL, Baird L, Seravalli J, Dien BS, Sarath G, Clemente TE, Sattler SE (2018) Overexpression of the *Sorghum bicolor* SbCCoAOMT alters cell wall associated hydroxycinnamoyl groups. PLoS ONE 13:e0204153

Towers GHN, Gibbs RD (1953) Lignin chemistry and the taxonomy of higher plants. Nature 172:25–26

Vanholme R, Storme V, Vanholme B, Sundin L, Christensen JH, Goeminne G, Halpin C, Rohde A, Morreel K, Boerjan W (2012) A systems biology view of responses to lignin biosynthesis perturbations in *Arabidopsis*. Plant Cell. https://doi.org/10.1105/tpc.112.102574

Vermerris W, Thompson KJ, Mcintyre LM, Axtell JD (2002) Evidence for an evolutionarily conserved interaction between cell wall biosynthesis and flowering in maize and sorghum. BMC Evol Biol 2:2

Vignols F, Rigau J, Torres MA, Capellades M, Puigdomenech P (1995) The brown midrib3 (bm3) mutation in maize occurs in the gene encoding caffeic acid *O*-methyltransferase. Plant Cell 7:407–416

Vogel KP, Pedersen JF, Masterson SD, Toy JJ (1999) Evaluation of a filter bag system for NDF, ADF, and IVDMD forage analysis. Crop Sci 39:276–279

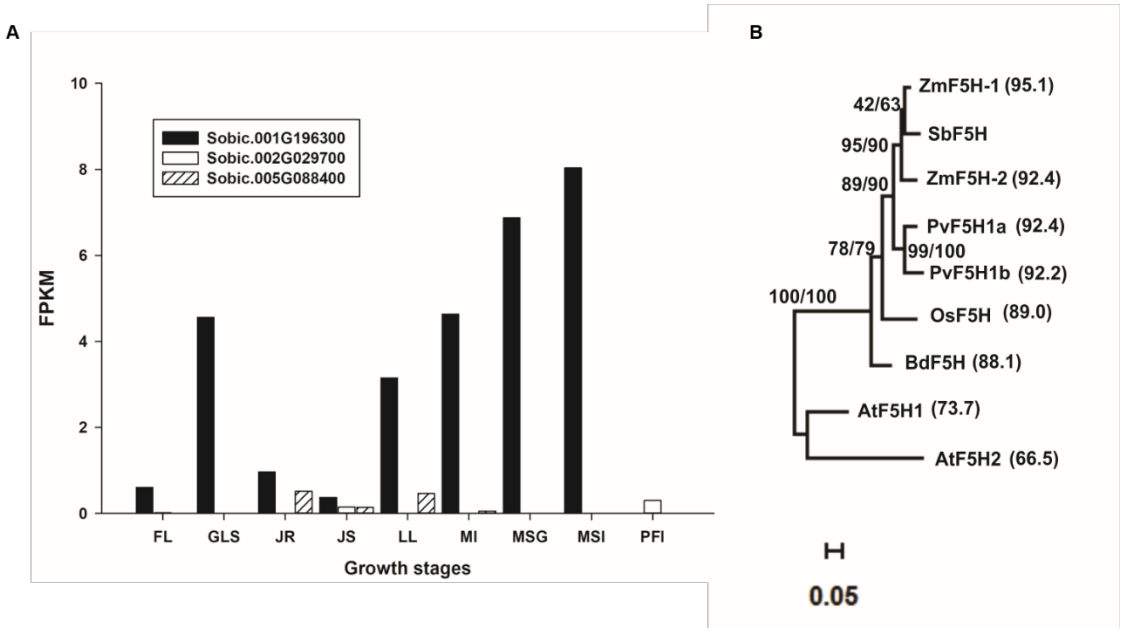
Weng JK, Mo H, Chapple C (2010) Over-expression of F5H in COMT-deficient Arabidopsis leads to enrichment of an unusual lignin and disruption of pollen wall formation. Plant J 64:898–911

Wu Z, Wang N, Hisano H, Cao Y, Wu F, Liu W, Bao Y, Wang Z-Y, Fu C (2019) Simultaneous regulation of F5H in COMT-RNAi transgenic switchgrass alters effects of COMT suppression on syringyl lignin biosynthesis. Plant Biotechnol J 17:836–845

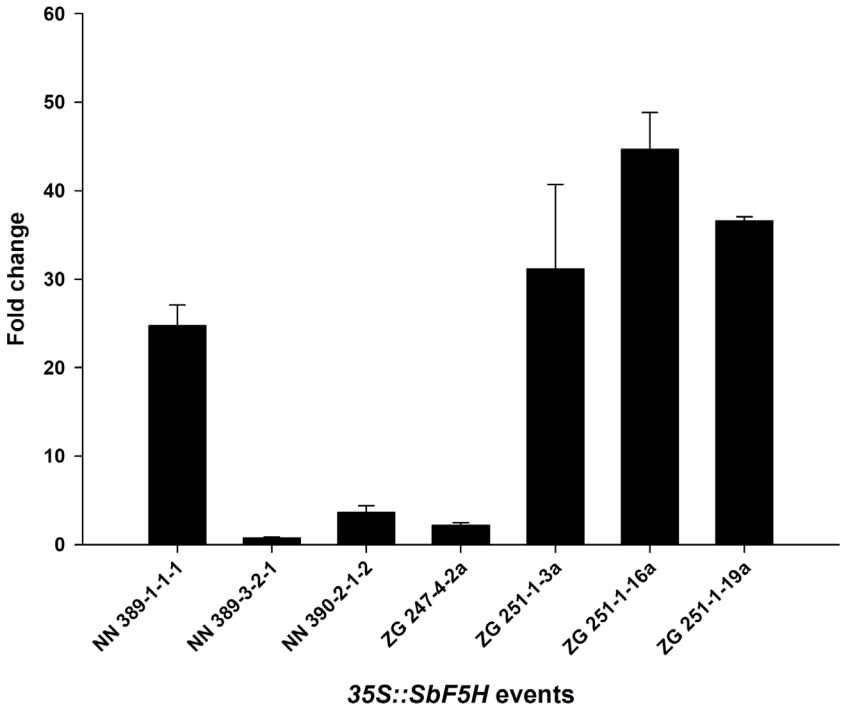
Yue F, Lu F, Sun R-C, Ralph J (2012) Syntheses of lignin-derived thioacidolysis monomers and their uses as quantitation standards. J Agric Food Chem 60:922–928

Yue F, Lu F, Regner M, Sun R, Ralph J (2017) Lignin-derived thioacidolysis dimers: reevaluation, new products, authentication, and quantification. Chemsuschem 10:830–835

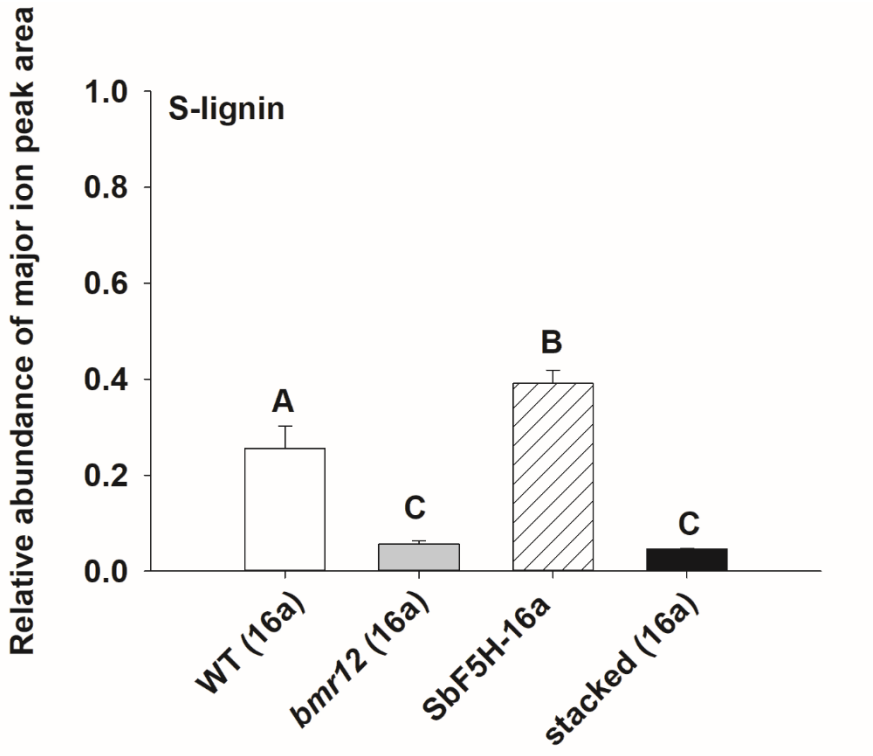
Zhao H, Wei J, Jinyu Z, Liu H, Tai W, Song Y (2002) Lignin biosynthesis by suppression of two O-methyl-transferases. Chin Sci Bull 47:1092–1095



Supplemental Fig. 1 Characterization of *SbF5H* (Sobic.001G196300). A) Expression of *Sorghum bicolor F5H* (Sobic.001G196300 (*SbF5H*), Sobic.002G029700 and Sobic.005G088400) for various tissues throughout development from Phytozome (https://phytozome.jgi.doe.gov/pz/portal.html). **B)** Phylogenetic analysis of *F5H* in sorghum, Arabidopsis, maize, brachypodium, rice and switchgrass based on amino acid sequences. A maximum likelihood tree depicting F5H based on 494 amino acid residues. Numbers along branches indicate bootstrap support for Maximum Likelihood/Neighbor-joining analyses. Numbers in parentheses represent percent sequence similarity to *SbF5H*. Sequence data was obtained from Phytozome with the following accession numbers: ZmF5H-1 (AC210173.4), SbF5H (Sobic.001G196300), ZmF5H-2 (GRMZM2G100158), PvF5H1a (Pavir.Ia01427), PvF5H1b (Pavir.Ib03743), OsF5H (LOC Os10g36848), BdF5H (Bradi3g30590), AtF5H1 (AT4G46220) and AtF5H2 (AT5G04330). FPKM, Fragments Per Kilobase of transcripts per Million mapped reads; FL, flag leaf; GLS, growing leaf sheath; JR, juvenile root; JS, juvenile stem; LL, lower leaf; MI, mature internode under peduncle; MSG, mature seed grain; MSI, mature stem internode; PFI, panicle floral initiation.



Supplemental Fig. 2 Gene expression for *35S::SbF5H* in the seven independent transformation events. NN and ZG are abbreviations for the technical staff that conducted transformation experiments. Expression was determined using the ΔΔCt method with the α-tubulin gene (Sobic.001G1070200.1) and wild-type for normalization. Values presented are means; SE represent standard error. Two biological replicates were used for analysis.



Supplemental Fig 3. The abundance of major ion peak area of syringyl (S-lignin) subunit composition determined using thioacidolysis for wild-type (WT), wild-type, *bmr12*, *35S::SbF5H* and stacked (*bmr12 35S::SbF5H*) stover relative to the internal standard (4,4'-Ethylidenebisphenol) were determined by GC/MS ($p < 0.0001$). Values presented represent least square means; error bars represent standard error. Samples with different letters are statistically different at $\alpha \leq 0.05$ using Tukey's HSD test.

Supplemental Table S1. Primers used for RT-qPCR.

Supplemental Table S2. Relative expression levels of genes from the monolignol biosynthetic pathway determined using BioMark HD, a microfluidic-based, high throughput qPCR system. Relative expression was determined using the ΔCt method with the α-tubulin gene (Sobic.001G1070200.1) for normalization. Values presented are means; SE represent standard error. PAL, phenylalanine ammonia lyase; C4H, cinnamate 4-hydroxylase; 4CL, 4-coumarate-CoA ligase: HCT, *p*-hydroxycinnamoyltransferase; C3'H, *p*-coumaroyl quinate/shikimate 3'-hydroxylase; CCoAOMT, caffeoyl-CoA-O-methyltransferase; CCR, cinnamoyl-CoA reductase; F5H, ferulate 5-hydroxylase; COMT, caffeic acid O-methyltransferase; CAD, cinnamyl alcohol dehydrogenase; DAHP, 3-deoxy-o-arabino-heptulosonate; FPGS, folypolyglutamate synthase; MTHFR, methylenetetrahydrofolate reductase.

Supplemental Table S3. Analysis of soluble and wall-bound phenolics from wild-type, *bmr12*, *35S::SbF5H* and stacked (*bmr12 35S::SbF5H*) plants. Phenolic acids were analyzed via GC-MS. Values presented are least square means; SE represent standard error. Samples with different letters for wall-bound and soluble fractions are statistically different from one another at $\alpha \leq 0.05$ using Tukey's HSD test.